

FASTA searches a protein or DNA sequence data bank  
version 3.3t05 March 30, 2000

Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

/tmp/fastaCAAQPaiHi: 995 aa  
>SEQ ID NO:2  
vs /tmp/fastaDAARPaiHi library  
searching /tmp/fastaDAARPaiHi library

1008 residues in 1 sequences

FASTA (3.34 January 2000) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 39, opt: 27, gap-pen: -12/-2, width: 16  
Scan time: 0.050

The best scores are: opt  
NM\_000096 ACCESSION:NM\_000096 NID: gi 4557484 ref (1008) 2671

>>NM\_000096 ACCESSION:NM\_000096 NID: gi 4557484 ref NM\_0 (1008 aa)  
initn: 1414 initl: 972 opt: 2671  
Smith-Waterman score: 3980; 57.002% identity in 1014 aa overlap (1-992:1-1008)

```

      10      20      30      40      50
SEQ   MKILILGIFLFLCSPGWAIDRHCYIGIEESIWNYPSPGKNMLNEKPFSEDL-----FLQ
      .....:.....:.....:.....:.....:.....:.....:.....:.....:.....:
NM_000 MKILILGIFLFLCSTPAWAKEKHYYIGIETTWDYA---SDHGEKKLISVDTEHSNIYLO
      10      20      30      40      50
      60      70      80      90     100     110
SEQ   GGQARKSFVFKKALYFYTDNTFQRIIEKPSWLGLGPMIKAETGDFIYVHVKNNASRAY
      ..:.....:.....:.....:.....:.....:.....:.....:.....:.....:
NM_000 NGPDRIGRLYKKALYLQYTDFTRTTIEKPVWLGLGPIIKAETGDKVYVHLKNLASRPY
      60      70      80      90     100     110
      120     130     140     150     160     170
SEQ   SYHPHGLTYSKENEAHGAIYPDNTTGLQKEVEYLEPGKQYTYKWYVEEHQPGPNDSNCV
      ...:.....:.....:.....:.....:.....:.....:.....:.....:.....:
NM_000 TFHSHGITYYKEHE--GAIYPDNTTDFQRADDKVYPGEQYTYMLLATEEQSPGEGDGNCV
      120     130     140     150     160     170
      180     190     200     210     220     230
SEQ   TRIYHSHIDTARDVASGLIGPILTCRGTNLNGDTEKDIDRSSFLMFSTTDESRSWYSDEN
      .....:.....:.....:.....:.....:.....:.....:.....:.....:.....:
NM_000 TRIYHSHIDAPKDIASGLIGPLIICKKDSLDKEKEKHIDREFVVMFSVVDENFSWYLEDN
      180     190     200     210     220     230
      240     250     260     270     280     290
SEQ   IRAF-TESGKINTSDPRFEESMSMQSINGYIYGNLPNLTMCAEDRVQWYFVGMGGVADIH
      ....:.....:.....:.....:.....:.....:.....:.....:.....:.....:
NM_000 IKTYCSEPEKVDKDNEFDQESNRMYSVNGYTFGSLPGLSMCAEDRVKWYLFMGNEVDVH
      240     250     260     270     280     290
      300     310     320     330     340     350
SEQ   PVYLRGQTLISRNHRKDTIMLFPSSEDAFMVAKAPGVWMLGCQ-----IHESMQAFFFKVS
      .....:.....:.....:.....:.....:.....:.....:.....:.....:.....:
NM_000 AAFPHGQALTNKNYRIDTINLFPATLFDAYMVAQNPGEWMLSCQNLNHLKAGLQAFFQVQ
      300     310     320     330     340     350
```

	360	370	380	390	400	410
SEQ	NCQKPSTEAFVTGTHVIHYIIAAKEILWNYAPSGIDFFTCKNLTAAGSKSQLFFERSPTR					
NM_000	ECNKSSSKDNIRGKHVRHYIIAAEELIWNYPGIDIFTKENLTAPGSDSAVFEEQGTTR					
	360	370	380	390	400	410
	420	430	440	450	460	
SEQ	IGGTYKKLIYREYTDASFQTQKAR---EEHLGILGPVFKAIEVGQTITIKITFYNNASLPLSI					
NM_000	IGGSYKKLVYREYTDASFNTNRKERGPREEHLGILGPVIWAEVGDTIRVTFHNGKAYPLSI					
	420	430	440	450	460	470
	470	480	490	500	510	520
SEQ	QPPGLHYNKSNEGLFYE---TPGG-STPPSSSHVSPGTTFFVYTWEVPKDVGPTSTDPNCL					
NM_000	EPIGVRFNKNNEGTYYSPNYPQSRVPPSASHVAPTETFTYEWTVPKEVGPTNADPVCL					
	480	490	500	510	520	530
	530	540	550	560	570	580
SEQ	TWFFYYSSVNGKKDINSGLLGPLLICRNGSLGDDGKQKGVDFEYLLATIFDENESNLLDE					
NM_000	AKMYYSADVPTKDIFTGLIGPMKICKKGS LHANGRQKDVDKEFYLFPTVFDENESLLED					
	540	550	560	570	580	590
	590	600	610	620	630	640
SEQ	N-RTFITEPENIDKEDTDCQASNKMYSSINGYMYGNLPGLDTC LGDNVLWHVFSVGSVEDL					
NM_000	NIRMFTTAPDQVDKEDEDFQESNKMHSNMGFMYGNQPLTMCKGDSVVWYLF SAGNEADV					
	600	610	620	630	640	650
	650	660	670	680	690	700
SEQ	HGIYFSGNTFTSLGARRDTIPMPFYTSQTLLMTPDSIGTFDLVCMTIKHN LGGMKHKYHV					
NM_000	HGIYFSGNTYLWRGERRD TANLFPQTS LTLHMWPDTEGTFNVECLTDDHYTGGMKQKYTV					
	660	670	680	690	700	710
	710	720	730	740	750	760
SEQ	RQCGKPNPDQTQYQEEKIIITIAAEEMEW DYSPSRKWENELHHLRRENQTS MYVDRSGTL					
NM_000	NQCRRQSEDSTFYLGERTYY-IAAVEVEWDYSPQREWEKELHHLQEQNVSN AFLDKGEFY					
	720	730	740	750	760	770
	770	780	790	800	810	820
SEQ	LGSKYKKVLYRQYDDNTFTNQTKRNEGEKHL DILGPLILLNPGQIIQIIFKNKAARPYSI					
NM_000	IGSKYKKVVYRQYTDSTFRVPVERKAE EHLGILGPQLHADVGDKVKIIFKNMATRPYSI					
	780	790	800	810	820	830
	830	840	850	860	870	880
SEQ	HAHGVTNNSTVVP TQPGEIQIYTWQIPDR TGPTSLDFECIPWFYYSTVSVAKDLHSGLV					
NM_000	HAHGVTESSTVTP TLPGETLYVWKIPERSGAGTEDSACIPWAYYSTVDQVKDLYSGLI					
	840	850	860	870	880	890

```

      890      900      910      920      930
SEQ      GPLSVCR----KDINPN-IVHRVLHFMIFDENESWYFEDSINTYASKPNKVDKENDNFQL
      ::: :::      : :::      .. : : : : : : : : : : : : : : : : : : : : : :
NM_000    GPLIVCRPYLKVFNP RRKLEFALLFLVFDENESWYLD DNIKTYS DHPKVNKDDEEFIE
      900      910      920      930      940      950

      940      950      960      970      980      990
SEQ      SNQMHAINGR LFGNNQGITFHVGDVVNWYLIGIGNEADLHTVHFHGHGSFEYKHKYLI
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
NM_000    SNKMHAINGRMFGNLQGLTMHVGDEVNWYLMGMGNEIDLHTVHFHGHGSFQYKHR
      960      970      980      990      1000

```

Function used was FASTA